Statistics 1040, Section 006, Quiz 1 (20 Points)

Friday, September 3, 2004

Your Name: _____

Question 1: Controlled Experiments/Observational Studies I (**14 Points**)

(*hypothetical*) Does regularly taking vitamin C help protect people against flu?

A _____ was conducted to answer this question. The _____ were 500 volunteering college students, assigned _____ to two groups of 250 students. The students in the _____ took regularly a tablet of vitamin C, whereas those in the _____ took an identically looking and tasting pill, called _____. Neither participating students nor personell administrating drugs to them knew who was taking which pill, in other words, it was a _____ experiment. After a couple of months, the numbers of flu cases in both groups were compared ...

Fill the gaps in the paragraph above using the most appropriate words from the following list:

placebo
double–blind
haphazardly
treatment group
observational study
randomly
single–blind
vaccine
confounding factor
objects
control group
controlled experiment
subjects
polio

Please turn over!

Question 2: Controlled Experiments/Observational Studies II (**6 Points**)

In 1990, four passengers were killed by crashes on commuter airlines, compared to 39 killed on scheduled carriers (like United, TWA, and so forth). **True** or **false**? Circle your answer and explain: the data show that if you have to fly, it is safer to do so on a commuter airline.

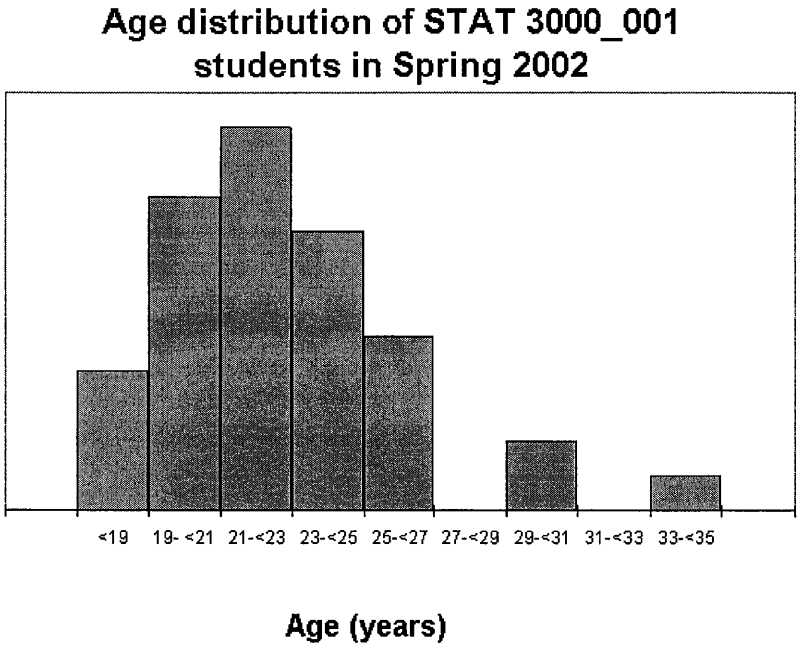
Statistics 1040, Section 006, Quiz 2 (20 Points)

Friday, September 10, 2004

Your Name: _____

Question 1: Histograms I (**12 Points**)

The histogram below shows the age distribution of Stat 3000, Section 001, students for the Spring 2002 semester. Unfortunately, the labels on the vertical axis have been deleted. However, the instructor recalls that there have been about 25% of students who were at least 21 but less than 23 years old. Try to help the instructor to fill in some of the missing percentages.



1. What approximate percentage of students were at least 25 but less than 27 years old?

2. What approximate percentage of students were younger than 21 years of age?

3. What approximate percentage of students were at least 29 years old?

Please turn over!

Question 2: Histograms II (**8 Points**)

An investigator draws a histogram for some height data, using the metric system. She is working in centimeters (cm). The vertical axis shows density, and the top of the vertical axis is 10 percent per cm. Now she wants to convert to millimeters (mm). There are 10 millimeters to the centimeter. On the horizontal axis, she has to change 175 cm to _____ mm, and 200 cm to _____ mm. On the vertical axis, she has to change 10 percent per cm to _____ percent per mm, and 5 percent per cm to _____ percent per mm.

Statistics 1040, Section 006, Quiz 3 (20 Points)

Friday, September 17, 2004

Your Name: _____

Question 1: Measures of Center and Spread I (**10 Points**)

Suppose an advertisement reported that the average weight loss after using a certain exercise machine for 2 months was 10 pounds. You investigate further and discover that the median weight loss was 3 pounds.

1. Is it most likely that the histogram of all weight losses (i) has a long right tail, (ii) has a long left tail, or (iii) is symmetric? Circle your answer and explain!

2. As a **consumer** trying to decide whether to buy this exercise machine, would it be more helpful for you to know the mean or the median? Circle your answer and explain!

3. As a **sales manager of the company** producing these exercise machines, would it be more commercially profitable for you (and your company) to publish the mean or the median? Circle your answer and explain!

Please turn over!

Question 2: Measures of Center and Spread II (**10 Points**)

Find the average and the standard deviation of the following two lists of numbers:

	Numbers	Average	Standard deviation
List 1:	7, 7, 7, 7	_____	_____
List 2:	7, –7, 7, –7	_____	_____

Show your work and/or give a short explanation for your answer.
Note: A calculation is NOT always necessary!
Use the formulas provided below (if needed).

Formulas:

$$\text{avg} = \frac{\text{sum of all numbers}}{\text{how many numbers}}$$

$$\text{SD} = \sqrt{\text{average of } [(\text{deviations from avg})^2]}$$

Statistics 1040, Section 006, Quiz 4 (20 Points)

Friday, September 24, 2004

Your Name: _____

Question 1: Normal Approximation for Data (**20 Points**)

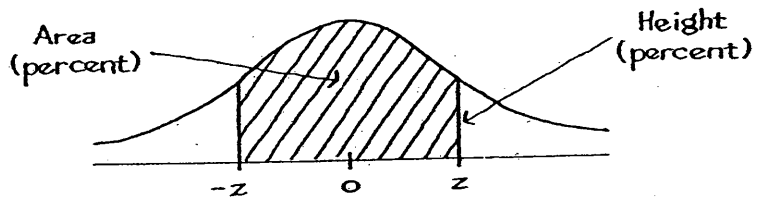
The heights of women approximately follow the normal curve with an average of **65 inches** and a standard deviation of **2.0 inches**. Answer the questions below:

- (**12 Points**) The percentage of women who are between 62 and 70 inches tall is roughly _____ %.

- (**8 Points**)
  The percentage of of women who are taller than 72 inches is about _____ %.

Show your work!

Tables



A NORMAL TABLE

z	*Area*	*z*	*Area*	*z*	*Area*
0.00	0	1.50	86.64	3.00	99.730
0.05	3.99	1.55	87.89	3.05	99.771
0.10	7.97	1.60	89.04	3.10	99.806
0.15	11.92	1.65	90.11	3.15	99.837
0.20	15.85	1.70	91.09	3.20	99.863
0.25	19.74	1.75	91.99	3.25	99.885
0.30	23.58	1.80	92.81	3.30	99.903
0.35	27.37	1.85	93.57	3.35	99.919
0.40	31.08	1.90	94.26	3.40	99.933
0.45	34.73	1.95	94.88	3.45	99.944
0.50	38.29	2.00	95.45	3.50	99.953
0.55	41.77	2.05	95.96	3.55	99.961
0.60	45.15	2.10	96.43	3.60	99.968
0.65	48.43	2.15	96.84	3.65	99.974
0.70	51.61	2.20	97.22	3.70	99.978
0.75	54.67	2.25	97.56	3.75	99.982
0.80	57.63	2.30	97.86	3.80	99.986
0.85	60.47	2.35	98.12	3.85	99.988
0.90	63.19	2.40	98.36	3.90	99.990
0.95	65.79	2.45	98.57	3.95	99.992
1.00	68.27	2.50	98.76	4.00	99.9937
1.05	70.63	2.55	98.92	4.05	99.9949
1.10	72.87	2.60	99.07	4.10	99.9959
1.15	74.99	2.65	99.20	4.15	99.9967
1.20	76.99	2.70	99.31	4.20	99.9973
1.25	78.87	2.75	99.40	4.25	99.9979
1.30	80.64	2.80	99.49	4.30	99.9983
1.35	82.30	2.85	99.56	4.35	99.9986
1.40	83.85	2.90	99.63	4.40	99.9989
1.45	85.29	2.95	99.68	4.45	99.9991

Statistics 1040, Section 006, Quiz 5 (20 Points)

Friday, October 1, 2004

Your Name: _____

Question 1: Percentiles and the Normal Curve (**12 Points**)

Among freshmen at a certain university, scores on the Math SAT followed the normal curve, with an average of 500 and an SD of 100. Fill in the blanks. **Show your work!**

1. (**6 Points**) A student who scored 350 on the Math SAT was at the _____th percentile of the score distribution.

2. (**6 Points**) To be at the 75th percentile of the distribution, a student needed a score of about _____ points on the Math SAT.

Question 2: Correlation (**8 Points**)

Investigators take a sample of DINKS (dual–income families, where husband and wife both work and have no kids). The investigators have data on the husband's income and the wife's income. By definition,

family income = husband's income + wife's income.

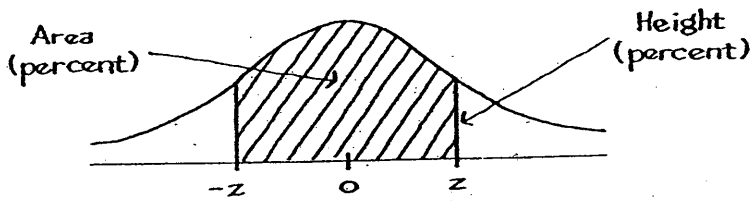
The average family income was around \$50,000, and 10% of the couples had family income in the range \$45,000–\$55,000. Fill in the blanks, using the options given below, and **explain briefly**:

1. (**4 Points**) The correlation between wife's income and family income is _____.

2. (**4 Points**) Among couples whose family income is in the range \$45,000–\$55,000, the correlation between wife's income and husband's income is _____.

Options: (a) -1 (b) nearly -1 (c) somewhat negative (d) 0
(e) somewhat positive (f) nearly 1 (g) 1

Tables



A NORMAL TABLE

z	Area	z	Area	z	Area
0.00	0	1.50	86.64	3.00	99.730
0.05	3.99	1.55	87.89	3.05	99.771
0.10	7.97	1.60	89.04	3.10	99.806
0.15	11.92	1.65	90.11	3.15	99.837
0.20	15.85	1.70	91.09	3.20	99.863
0.25	19.74	1.75	91.99	3.25	99.885
0.30	23.58	1.80	92.81	3.30	99.903
0.35	27.37	1.85	93.57	3.35	99.919
0.40	31.08	1.90	94.26	3.40	99.933
0.45	34.73	1.95	94.88	3.45	99.944
0.50	38.29	2.00	95.45	3.50	99.953
0.55	41.77	2.05	95.96	3.55	99.961
0.60	45.15	2.10	96.43	3.60	99.968
0.65	48.43	2.15	96.84	3.65	99.974
0.70	51.61	2.20	97.22	3.70	99.978
0.75	54.67	2.25	97.56	3.75	99.982
0.80	57.63	2.30	97.86	3.80	99.986
0.85	60.47	2.35	98.12	3.85	99.988
0.90	63.19	2.40	98.36	3.90	99.990
0.95	65.79	2.45	98.57	3.95	99.992
1.00	68.27	2.50	98.76	4.00	99.9937
1.05	70.63	2.55	98.92	4.05	99.9949
1.10	72.87	2.60	99.07	4.10	99.9959
1.15	74.99	2.65	99.20	4.15	99.9967
1.20	76.99	2.70	99.31	4.20	99.9973
1.25	78.87	2.75	99.40	4.25	99.9979
1.30	80.64	2.80	99.49	4.30	99.9983
1.35	82.30	2.85	99.56	4.35	99.9986
1.40	83.85	2.90	99.63	4.40	99.9989
1.45	85.29	2.95	99.68	4.45	99.9991

Statistics 1040, Section 006, Quiz 6 (20 Points)

Friday, October 15, 2004

Your Name: _____

Question 1: The Regression Line (**20 Points**)

A researcher is interested in the extent to which lead particles emitted from automobiles are absorbed by competitive cyclists. For a large group of cyclists they found the following:

Hours of training:	average = 16.2,	SD = 5.9	
Blood lead (μmol/L):	average = .42,	SD = .19,	r=0.6.

The scatter plot of the data is football–shaped.

Show your work!

1. (**7 Points**) Find the equation of the regression line for predicting blood lead from training time.

   Equation: _____

2. (**3 Points**) Use the regression equation from part 1. to predict the blood lead for a cyclist who trained for 21 hours.

   Answer: _____

Please turn over!

3. (**5 Points**) Find the r.m.s. error for predicting blood lead from training time of cyclist.

   Answer: _____

4. (**5 Points**) Would you be surprised to learn that a cyclist who trained for 3 hours had a blood lead of .8 μmol/L? Support your answer with a brief explanation and calculation.

   Answer: **Yes, surprised / No, not surprised**

   Explanation:

Formulas:

$$\text{r.m.s. error} = \sqrt{1 - r^2} \times \text{SD}_y$$

$$\text{slope} = r \times \frac{\text{SD}_y}{\text{SD}_x}$$

$$\text{intercept} = \text{avg}_y - \text{slope} \times \text{avg}_x$$

Statistics 1040, Section 006, Quiz 7 (20 Points)

Friday, October 22, 2004

Your Name: _____

Question 1: Chance/Probability (**20 Points**)

A drawer of socks contains 20 socks of which 6 are black, 10 are green, and 4 are blue. In the dark, a child chooses two socks at random to wear to school.

1. (**5 Points**) What is the chance that the first sock is blue?

   The chance is: _____ %

2. (**5 Points**) What is the chance that both socks are blue?

   The chance is: _____ %

3. (**5 Points**) What is the chance that one sock is blue and the other sock is green?

   The chance is: _____ %

4. (**5 Points**) What is the chance that both socks are the same color?

   The chance is: _____ %

Statistics 1040, Section 006, Quiz 8 (20 Points)

Friday, October 29, 2004

Your Name: _____

Question 1: Box Models, EV, and SE (**16 Points**)

You are participating in a new game that consists of tossing a 10–sided die, with sides numbered from 1 to 10. The die is fair, i.e., it has the same chance of landing on any side. Every time the die shows a number that is a multiple of 3 (i.e., 3, 6, or 9) you win \$3, otherwise you lose \$1, except when the die lands on 10, in which case you win (or lose) nothing (\$0). Assume you are tossing this die 200 times.

1. (**4 Points**) Find the box model.

2. (**6 Points**) Find the expected value of your gain/loss.

3. (**6 Points**) Find the standard error of your gain/loss.

Please turn over!

Question 2: Law of Averages (**4 Points**)

A box contains red and green marbles; there are more red marbles than green ones. Marbles are drawn one at a time from the box, at random with replacement. You win a dollar if a red marble is drawn more often than a green one. There are two choices:

- A: 50 draws are made from the box.
- B: 500 draws are made from the box.

Choose (i.e., circle) one of the four options below. **Explain your answer.**

1. A gives a better chance of winning.
2. B gives a better chance of winning.
3. A and B give the same chance of winning.
4. Can't tell without more information.

Formulas:

$$\text{box average} = \frac{\text{sum of all numbers in box}}{\text{how many numbers in box}}$$

$$\text{box SD} = \sqrt{\text{average of [(deviations from box average)}^2]}$$

$$\text{EV}_{sum} = \text{number of draws} \times \text{box average}$$

$$\text{SE}_{sum} = \sqrt{\text{number of draws}} \times \text{box SD}$$

Statistics 1040, Section 006, Quiz 9 (20 Points)

Friday, November 5, 2004

Your Name: _____

Question 1: EV, SE, and Normal Curve (**15 Points**)

In a certain town, there are 40,000 registered voters, of whom 15,000 are Democrats. A survey organization is about to take a simple random sample of 1,000 registered voters. **Show your work!**

1. (**5 Points**) Find the box model.

2. (**5 Points**) The expected number of Democrats in this sample of 1,000 is _____ with an SE of _____.

3. (**5 Points**) The chance that **at least** 500 of the voters in the sample are Democrats is about _____ %.

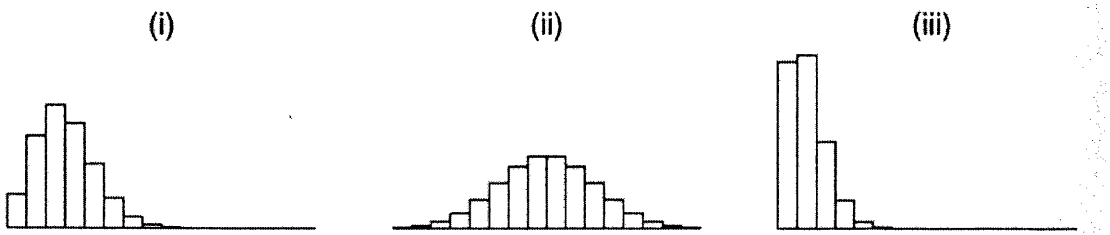
Please turn over!

Question 2: Probability Histograms (**5 Points**)

Twenty–five draws are made at random with replacement from the each of the boxes below:

A) $\boxed{\boxed{0}\ \boxed{1}}$ B) $\boxed{9\ \boxed{0}\text{ 's }\boxed{1}}$ C) $\boxed{24\ \boxed{0}\text{ 's }\boxed{1}}$

The probability histograms for the sums are shown below, in scrambled order. Match the histogram with the boxes. **Briefly explain your choices.**



(i) goes with _____; (ii) goes with _____; (iii) goes with _____

Explanation:

Formulas:

$$\text{box average} = \frac{\text{sum of all numbers in box}}{\text{how many numbers in box}}$$

$$\text{box SD} = \sqrt{\text{average of [(deviations from box average)}^2]}$$

$$\text{EV}_{sum} = \text{number of draws} \times \text{box average}$$

$$\text{SE}_{sum} = \sqrt{\text{number of draws}} \times \text{box SD}$$

Shortcut formulas for a box that contains only *two* different numbers:

$$\text{average} = \frac{(\text{smaller} \times \text{how many}) + (\text{bigger} \times \text{how many})}{\text{how many tickets in the box}}$$

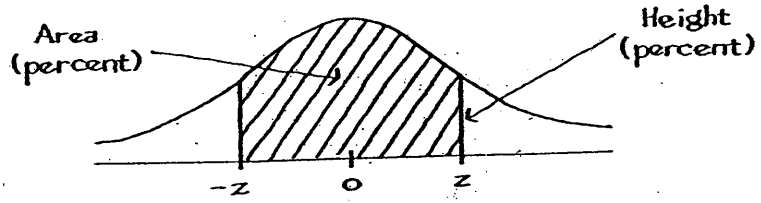
$$\text{SD} = (\text{bigger - smaller}) \times \sqrt{\frac{\text{fraction}}{\text{bigger}} \times \frac{\text{fraction}}{\text{smaller}}}$$

Shortcut formulas for a box that contains only $\boxed{0}$'s and $\boxed{1}$'s:

$$\text{average} = \frac{\text{number of } \boxed{1} \text{ 's}}{\text{how many tickets in the box}}$$

$$\text{SD} = \sqrt{\frac{\text{fraction}}{\text{of } \boxed{1} \text{ 's}} \times \frac{\text{fraction}}{\text{of } \boxed{0} \text{ 's}}}$$

Tables



A NORMAL TABLE

z	Area	z	Area	z	Area
0.00	0	1.50	86.64	3.00	99.730
0.05	3.99	1.55	87.89	3.05	99.771
0.10	7.97	1.60	89.04	3.10	99.806
0.15	11.92	1.65	90.11	3.15	99.837
0.20	15.85	1.70	91.09	3.20	99.863
0.25	19.74	1.75	91.99	3.25	99.885
0.30	23.58	1.80	92.81	3.30	99.903
0.35	27.37	1.85	93.57	3.35	99.919
0.40	31.08	1.90	94.26	3.40	99.933
0.45	34.73	1.95	94.88	3.45	99.944
0.50	38.29	2.00	95.45	3.50	99.953
0.55	41.77	2.05	95.96	3.55	99.961
0.60	45.15	2.10	96.43	3.60	99.968
0.65	48.43	2.15	96.84	3.65	99.974
0.70	51.61	2.20	97.22	3.70	99.978
0.75	54.67	2.25	97.56	3.75	99.982
0.80	57.63	2.30	97.86	3.80	99.986
0.85	60.47	2.35	98.12	3.85	99.988
0.90	63.19	2.40	98.36	3.90	99.990
0.95	65.79	2.45	98.57	3.95	99.992
1.00	68.27	2.50	98.76	4.00	99.9937
1.05	70.63	2.55	98.92	4.05	99.9949
1.10	72.87	2.60	99.07	4.10	99.9959
1.15	74.99	2.65	99.20	4.15	99.9967
1.20	76.99	2.70	99.31	4.20	99.9973
1.25	78.87	2.75	99.40	4.25	99.9979
1.30	80.64	2.80	99.49	4.30	99.9983
1.35	82.30	2.85	99.56	4.35	99.9986
1.40	83.85	2.90	99.63	4.40	99.9989
1.45	85.29	2.95	99.68	4.45	99.9991

Statistics 1040, Section 006, Quiz 10 (20 Points)

Friday, November 19, 2004

Your Name: _____

Question 1: Confidence Intervals (**20 Points**)

Recent political events are currently in focus of many surveys and polls nationwide. With four members of the Bush Cabinet resigning over the last few days, a natural concern for every U.S. citizen is: *Will the Bush Cabinet resignations have a positive or negative impact on U.S. policy?*

This question was asked to a sample of 787 U.S. citizens: 299 of them answered "Positive".

1. (**14 Points**) Construct a 87% confidence interval for the percentage of all U.S. citizens who think that the Bush Cabinet resignations will have a positive impact on U.S. policy.

   Show your work.

Please turn over!

2. (**6 Points**) For each of the following situations, explain **why** or **why not** it would be possible to constuct a 87% confidence interval for the percentage of all U.S. citizens who think that the Bush Cabinet resignations will have positive impact on U.S. policy. **Please do not construct the actual confidence interval – just answer each question with Yes or No and provide a very brief explanation.**

   - The sample of 787 U.S. citizens was obtained by using a computer to randomly generate a sufficent number of valid telephone numbers (including area code) and calling these numbers until 787 valid answers were collected.

     Is it possible to construct a 87% CI here? – **Yes** or **No**?

     Explanation:

   - The sample of 787 U.S. citizens was obtained as a SRS from all U. S. citizens, but 780 of the responders said "Positive" (i.e., thought that the Bush Cabinet resignations will have positive impact on U.S. policy).

     Is it possible to construct a 87% CI here? – **Yes** or **No**?

     Explanation:

   - The 787 answers come from the Quick Poll at the CNN Web page (http://www.cnn.com).

     Is it possible to construct a 87% CI here? – **Yes** or **No**?

     Explanation:

Formulas:

$$\text{box average} = \frac{\text{sum of all numbers in box}}{\text{how many numbers in box}}$$

$$\text{box SD} = \sqrt{\text{average of [(deviations from box average)}^2]}$$

$$\text{EV}_{sum} = \text{number of draws} \times \text{box average}$$

$$\text{SE}_{sum} = \sqrt{\text{number of draws}} \times \text{box SD}$$

Shortcut formulas for a box that contains only *two* different numbers:

$$\text{average} = \frac{(\text{smaller} \times \text{how many}) + (\text{bigger} \times \text{how many})}{\text{how many tickets in the box}}$$

$$\text{SD} = (\text{bigger - smaller}) \times \sqrt{\frac{\text{fraction}}{\text{bigger}} \times \frac{\text{fraction}}{\text{smaller}}}$$

Shortcut formulas for a box that contains only $\boxed{0}$'s and $\boxed{1}$'s:

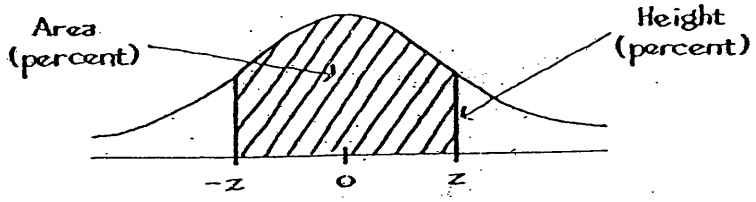
$$\text{average} = \frac{\text{number of } \boxed{1}\text{'s}}{\text{how many tickets in the box}}$$

$$\text{SD} = \sqrt{\frac{\text{fraction}}{\text{of } \boxed{1}\text{'s}} \times \frac{\text{fraction}}{\text{of } \boxed{0}\text{'s}}}$$

$$\text{EV}_{\%} = \%\text{ of } \boxed{1}\text{'s in the box}$$

$$\text{SE}_{\%} = \frac{\text{SE}_{\text{sum}}}{\#\text{draws}} \times 100\%$$

Tables



A NORMAL TABLE

z	Area	z	Area	z	Area
0.00	0	1.50	86.64	3.00	99.730
0.05	3.99	1.55	87.89	3.05	99.771
0.10	7.97	1.60	89.04	3.10	99.806
0.15	11.92	1.65	90.11	3.15	99.837
0.20	15.85	1.70	91.09	3.20	99.863
0.25	19.74	1.75	91.99	3.25	99.885
0.30	23.58	1.80	92.81	3.30	99.903
0.35	27.37	1.85	93.57	3.35	99.919
0.40	31.08	1.90	94.26	3.40	99.933
0.45	34.73	1.95	94.88	3.45	99.944
0.50	38.29	2.00	95.45	3.50	99.953
0.55	41.77	2.05	95.96	3.55	99.961
0.60	45.15	2.10	96.43	3.60	99.968
0.65	48.43	2.15	96.84	3.65	99.974
0.70	51.61	2.20	97.22	3.70	99.978
0.75	54.67	2.25	97.56	3.75	99.982
0.80	57.63	2.30	97.86	3.80	99.986
0.85	60.47	2.35	98.12	3.85	99.988
0.90	63.19	2.40	98.36	3.90	99.990
0.95	65.79	2.45	98.57	3.95	99.992
1.00	68.27	2.50	98.76	4.00	99.9937
1.05	70.63	2.55	98.92	4.05	99.9949
1.10	72.87	2.60	99.07	4.10	99.9959
1.15	74.99	2.65	99.20	4.15	99.9967
1.20	76.99	2.70	99.31	4.20	99.9973
1.25	78.87	2.75	99.40	4.25	99.9979
1.30	80.64	2.80	99.49	4.30	99.9983
1.35	82.30	2.85	99.56	4.35	99.9986
1.40	83.85	2.90	99.63	4.40	99.9989
1.45	85.29	2.95	99.68	4.45	99.9991

Statistics 1040, Section 006, Quiz 11 (20 Points)

Monday, November 29, 2004

Your Name: _____

Question 1: Confidence Intervals for Averages (**20 Points**)

When the latest Harry Potter book was released, a local bookstore sold 500 copies in one day. They randomly sampled 300 of these people and found that the average age for these 300 people was 13.3 years with a standard deviation of 8.2 years.

1. (**14 Points**) Find a 95% confidence interval for the average age of all 500 customers.

2. (**6 Points**) True or false? — Answer these three questions (**no** explanation needed):

   (a) 95% of the 300 sampled customers had ages in the interval from part 1.
   True or **False**? Just circle your answer.

   (b) We cannot approximate the percentage of customers who had ages in the interval from part 1. because the ages do not follow the normal curve.
   True or **False**? Just circle your answer.

   (c) The interval from part 1. is invalid because ages do not follow the normal curve.
   True or **False**? Just circle your answer.

Please turn over!

Formulas:

$$\text{box average} = \frac{\text{sum of all numbers in box}}{\text{how many numbers in box}}$$

$$\text{box SD} = \sqrt{\text{average of } [(\text{deviations from box average})^2]}$$

$$\text{EV}_{sum} = \text{number of draws} \times \text{box average}$$

$$\text{SE}_{sum} = \sqrt{\text{number of draws}} \times \text{box SD}$$

$$\text{EV}_{avg} = \text{box average} \qquad \text{SE}_{avg} = \frac{\text{SE}_{sum}}{\text{number of draws}}$$

Shortcut formulas for a box that contains only *two* different numbers:

$$\text{average} = \frac{(\text{smaller} \times \text{how many}) + (\text{bigger} \times \text{how many})}{\text{how many tickets in the box}}$$

$$\text{SD} = (\text{bigger - smaller}) \times \sqrt{\frac{\text{fraction}}{\text{bigger}} \times \frac{\text{fraction}}{\text{smaller}}}$$

Shortcut formulas for a box that contains only $\boxed{0}$'s and $\boxed{1}$'s:

$$\text{average} = \frac{\text{number of } \boxed{1}\text{'s}}{\text{how many tickets in the box}}$$

$$\text{SD} = \sqrt{\frac{\text{fraction}}{\text{of } \boxed{1}\text{'s}} \times \frac{\text{fraction}}{\text{of } \boxed{0}\text{'s}}}$$

$$\text{EV}_{\%} = \%\text{ of } \boxed{1}\text{'s in the box}$$

$$\text{SE}_{\%} = \frac{\text{SE}_{\text{sum}}}{\text{number of draws}} \times 100\%$$

Statistics 1040, Section 006, Quiz 12 (20 Points)

Friday, December 3, 2004

Your Name: _____

Question 1: Tests of Significance (**20 Points**)

Bookstores like education, because national data show that 71% of college graduates have read a book in the past year, compared to 54% of the general population age 18 and over. The data also show the nationwide average educational level to be 13 years of schooling completed, with an SD of about 3 years, for persons age 18 and over.
A bookstore is doing a market survey in a certain county, and takes a sample of 1,000 people age 18 and over. They find the average educational level to be 14 years, and the SD is 5 years. Can the difference in average educational level between the sample and the nation be explained by chance variation? If not, what other explanations can you give? Please **follow the steps below** in answering these questions.

1. (**5 Points**) State the null and the alternative hypothesis for this problem, in words and in terms of the box model.

2. (**5 Points**) Calculate the appropriate test statistic.

3. (**5 Points**) Obtain the P–value (use the normal table on the back).

4. (**5 Points**) State conclusions in terms of rejecting the null hypothesis and in your own words.

Please turn over!

Formulas:

$$\text{box average} = \frac{\text{sum of all numbers in box}}{\text{how many numbers in box}}$$

$$\text{box SD} = \sqrt{\text{average of } [(\text{deviations from box average})^2]}$$

$$\text{EV}_{sum} = \text{number of draws} \times \text{box average}$$

$$\text{SE}_{sum} = \sqrt{\text{number of draws}} \times \text{box SD}$$

$$\text{EV}_{avg} = \text{box average} \qquad \text{SE}_{avg} = \frac{\text{SE}_{sum}}{\text{number of draws}}$$

Shortcut formulas for a box that contains only *two* different numbers:

$$\text{average} = \frac{(\text{smaller} \times \text{how many}) + (\text{bigger} \times \text{how many})}{\text{how many tickets in the box}}$$

$$\text{SD} = (\text{bigger - smaller}) \times \sqrt{\frac{\text{fraction}}{\text{bigger}} \times \frac{\text{fraction}}{\text{smaller}}}$$

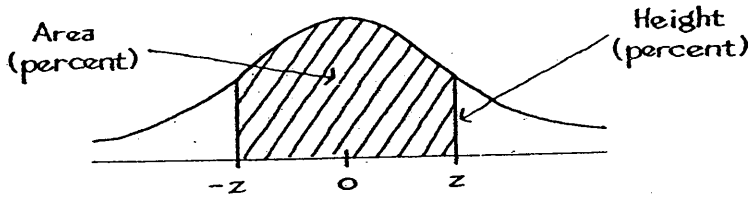
Shortcut formulas for a box that contains only $\boxed{0}$'s and $\boxed{1}$'s:

$$\text{average} = \frac{\text{number of } \boxed{1}\text{'s}}{\text{how many tickets in the box}}$$

$$\text{SD} = \sqrt{\frac{\text{fraction}}{\text{of } \boxed{1}\text{'s}} \times \frac{\text{fraction}}{\text{of } \boxed{0}\text{'s}}}$$

$$\text{EV}_{\%} = \%\text{ of } \boxed{1}\text{'s in the box} \qquad \text{SE}_{\%} = \frac{\text{SE}_{sum}}{\text{number of draws}} \times 100\%$$

Tables



A NORMAL TABLE

z	*Area*	*z*	*Area*	*z*	*Area*
0.00	0	1.50	86.64	3.00	99.730
0.05	3.99	1.55	87.89	3.05	99.771
0.10	7.97	1.60	89.04	3.10	99.806
0.15	11.92	1.65	90.11	3.15	99.837
0.20	15.85	1.70	91.09	3.20	99.863
0.25	19.74	1.75	91.99	3.25	99.885
0.30	23.58	1.80	92.81	3.30	99.903
0.35	27.37	1.85	93.57	3.35	99.919
0.40	31.08	1.90	94.26	3.40	99.933
0.45	34.73	1.95	94.88	3.45	99.944
0.50	38.29	2.00	95.45	3.50	99.953
0.55	41.77	2.05	95.96	3.55	99.961
0.60	45.15	2.10	96.43	3.60	99.968
0.65	48.43	2.15	96.84	3.65	99.974
0.70	51.61	2.20	97.22	3.70	99.978
0.75	54.67	2.25	97.56	3.75	99.982
0.80	57.63	2.30	97.86	3.80	99.986
0.85	60.47	2.35	98.12	3.85	99.988
0.90	63.19	2.40	98.36	3.90	99.990
0.95	65.79	2.45	98.57	3.95	99.992
1.00	68.27	2.50	98.76	4.00	99.9937
1.05	70.63	2.55	98.92	4.05	99.9949
1.10	72.87	2.60	99.07	4.10	99.9959
1.15	74.99	2.65	99.20	4.15	99.9967
1.20	76.99	2.70	99.31	4.20	99.9973
1.25	78.87	2.75	99.40	4.25	99.9979
1.30	80.64	2.80	99.49	4.30	99.9983
1.35	82.30	2.85	99.56	4.35	99.9986
1.40	83.85	2.90	99.63	4.40	99.9989
1.45	85.29	2.95	99.68	4.45	99.9991